TUGAS MINGGU KE-12

STATISTIKA DESKRIPTIF

NAMA : MUKHAMAD IKHSANUDIN

NIM : 082011633086

PROGRAM STUDI S1 SISTEM INFORMASI

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS AIRLANGGA

2021

Tugas pertemuan 24 → dikumpulkan hari ini, tgl. 04-06-2021 jam 23.59 → di upload ke Aula dan kirim ke email eto-w@fst.unair .ac.id dengan subject : PCA

Code dan outputnya jadikan satu di notebook R-nya
Gunakan prcomp() and princomp() functions untuk masing-masing soal berikut :

1. Carilah data yang sesuai untuk PCA dengan covarians
   a. Tampilkan eigenvalue-nya (variansnya) → table dan grafik
   b. Tampilkan matriks PC-nya
   c. Plot di 2-dimensi untuk individunya
   d. Plot di 2-dimensi untuk variabelnya
   e. Bi-Plot di 2-dimensi untuk individu dan variabelnya
   f. Plot PC di 3-dimensi
   g. Interpretasikan point c s.d f

2. Carilah data yang sesuai untuk PCA dengan correlation
   a. Tampilkan eigenvalue-nya (variansnya) → table dan grafik
   b. Tampilkan matriks PC-nya
   c. Plot di 2-dimensi untuk individunya
   d. Plot di 2-dimensi untuk variabelnya
   e. Bi-Plot di 2-dimensi untuk individu dan variabelnya
   f. Plot PC di 3-dimensi
   g. Interpretasikan point c s.d f

==========

Code ditaruh diantara tanda berikut :

```{R}

Syntax di sini

```

1. Carilah data yang sesuai untuk PCA dengan covarians

```r
#No 1
library(dplyr)
PCA_Data <- select(mtcars, c(1:7, 10:11))
head(PCA_Data[,-9])
```

Description: df[,8] [6 x 8]

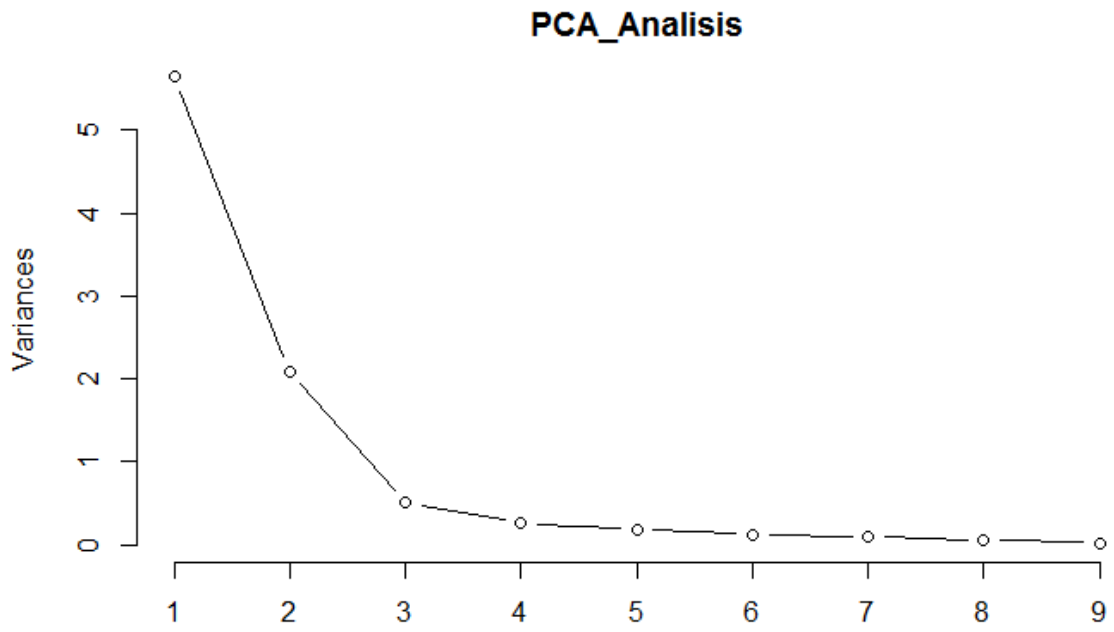
	mpg <dbl>	cyl <dbl>	disp <dbl>	hp <dbl>	drat <dbl>	wt <dbl>	qsec <dbl>	gear <dbl>
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	4
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	3
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	3
Valiant	18.1	6	225	105	2.76	3.460	20.22	3

6 rows

a. Tampilkan eigenvalue-nya (variansnya) → table dan grafik

```r
#a. Tampilkan eigen valuenya (variansnya)
PCA_Standardized <- scale(x = PCA_Data)
Covarian_Matrix <- cov(PCA_Standardized)
Nilai_Eigen <- eigen(Covarian_Matrix)
Nilai_Eigen
PCA_Analisis <- prcomp(x = PCA_Data, scale. = TRUE, center = TRUE)
PCA_Analisis
plot(PCA_Analisis, type = "l")
```

```
eigen() decomposition
$values
[1] 5.65593947 2.08210029 0.50421482 0.26502753 0.18315864 0.12379319 0.10506192 0.05851375
[9] 0.02219038

$vectors
           [,1]        [,2]        [,3]         [,4]       [,5]        [,6]        [,7]
 [1,]  0.3931477 -0.02753861 -0.22119309 -0.006126378 -0.3207620  0.72015586  0.38138068
 [2,] -0.4025537 -0.01570975 -0.25231615  0.040700251  0.1171397  0.22432550  0.15893251
 [3,] -0.3973528  0.08888469 -0.07825139  0.339493732 -0.4867849 -0.01967516  0.18233095
 [4,] -0.3670814 -0.26941371 -0.01721159  0.068300993 -0.2947317  0.35394225 -0.69620751
 [5,]  0.3118165 -0.34165268  0.14995507  0.845658485  0.1619259 -0.01536794 -0.04767957
 [6,] -0.3734771  0.17194306  0.45373418  0.191260029 -0.1874822 -0.08377237  0.42777608
 [7,]  0.2243508  0.48404435  0.62812782 -0.030329127 -0.1482495  0.25752940 -0.27622581
 [8,]  0.2094749 -0.55078264  0.20658376 -0.282381831 -0.5624860 -0.32298239  0.08555707
 [9,] -0.2445807 -0.48431310  0.46412069 -0.214492216  0.3997820  0.35706914  0.20604210
           [,8]        [,9]
 [1,]  0.12465987 -0.11492862
 [2,] -0.81032177 -0.16266295
 [3,]  0.06416707  0.66190812
 [4,]  0.16573993 -0.25177306
 [5,] -0.13505066 -0.03809096
 [6,]  0.19839375 -0.56918844
 [7,] -0.35613350  0.16873731
 [8,] -0.31636479 -0.04719694
 [9,]  0.10832772  0.32045892
```



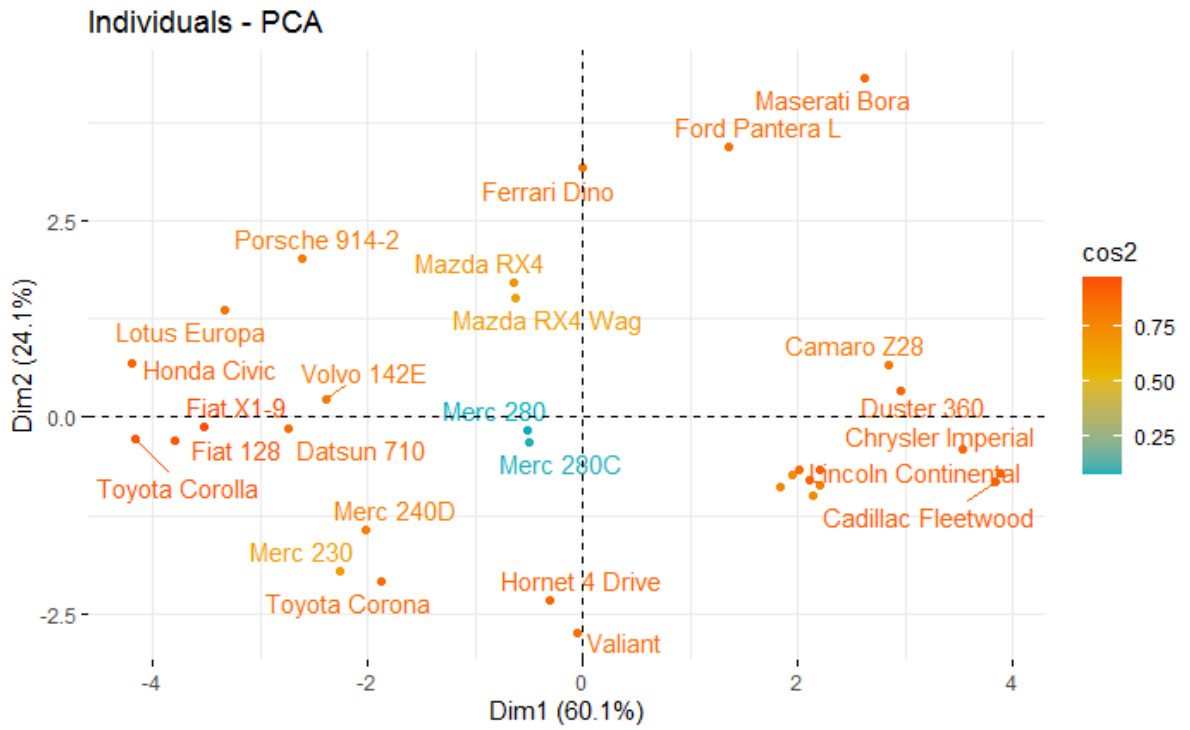
b. Tampilkan matriks PC-nya

```r
#b. tampilkan matriks PC nya
PCA_Analisis_New <- princomp(x = PCA_Data)
PCA_Analisis_New
Covarian_Matrix <- cov(PCA_Standardized)
Covarian_Matrix
```

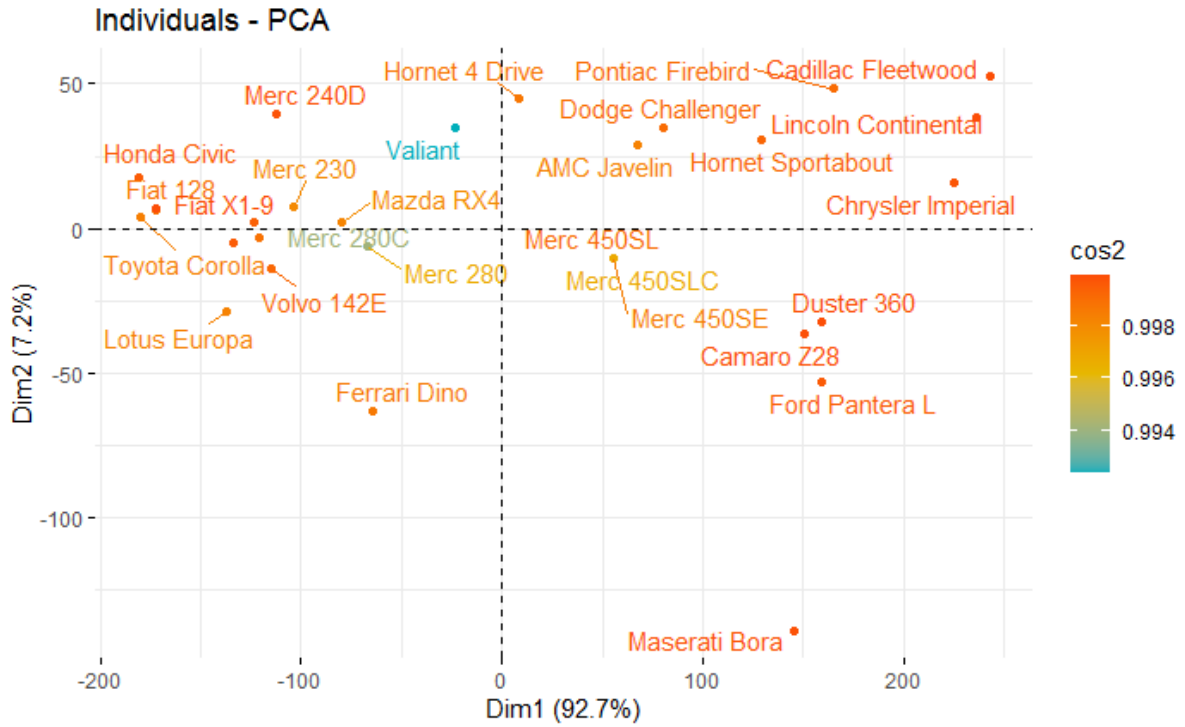
```
Call:
princomp(x = PCA_Data)

Standard deviations:
     Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6      Comp.7      Comp.8
134.3820274  37.5465620   3.0181296   1.2548450   0.8904901   0.6371404   0.3006062   0.2814188
     Comp.9
  0.2124807

 9  variables and  32 observations.
            mpg        cyl       disp         hp       drat         wt       qsec       gear
mpg   1.0000000 -0.8521620 -0.8475514 -0.7761684  0.68117191 -0.8676594  0.41868403  0.4802848
cyl  -0.8521620  1.0000000  0.9020329  0.8324475 -0.69993811  0.7824958 -0.59124207 -0.4926866
disp -0.8475514  0.9020329  1.0000000  0.7909486 -0.71021393  0.8879799 -0.43369788 -0.5555692
hp   -0.7761684  0.8324475  0.7909486  1.0000000 -0.44875912  0.6587479 -0.70822339 -0.1257043
drat  0.6811719 -0.6999381 -0.7102139 -0.4487591  1.00000000 -0.7124406  0.09120476  0.6996101
wt   -0.8676594  0.7824958  0.8879799  0.6587479 -0.71244065  1.0000000 -0.17471588 -0.5832870
qsec  0.4186840 -0.5912421 -0.4336979 -0.7082234  0.09120476 -0.1747159  1.00000000 -0.2126822
gear  0.4802848 -0.4926866 -0.5555692 -0.1257043  0.69961013 -0.5832870 -0.21268223  1.0000000
carb -0.5509251  0.5269883  0.3949769  0.7498125 -0.09078980  0.4276059 -0.65624923  0.2740728
           carb
mpg  -0.5509251
cyl   0.5269883
disp  0.3949769
hp    0.7498125
drat -0.0907898
wt    0.4276059
qsec -0.6562492
gear  0.2740728
carb  1.0000000
```

c. Plot di 2-dimensi untuk individunya

```{r}
#c. Plot di 2-dimensi untuk individunya
library(factoextra)
Visual_PCA_Cor <- prcomp(mtcars, scale = TRUE)
fviz_pca_ind(Visual_PCA_Cor, col.ind = "cos2",gradient.cols =
c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE)

```

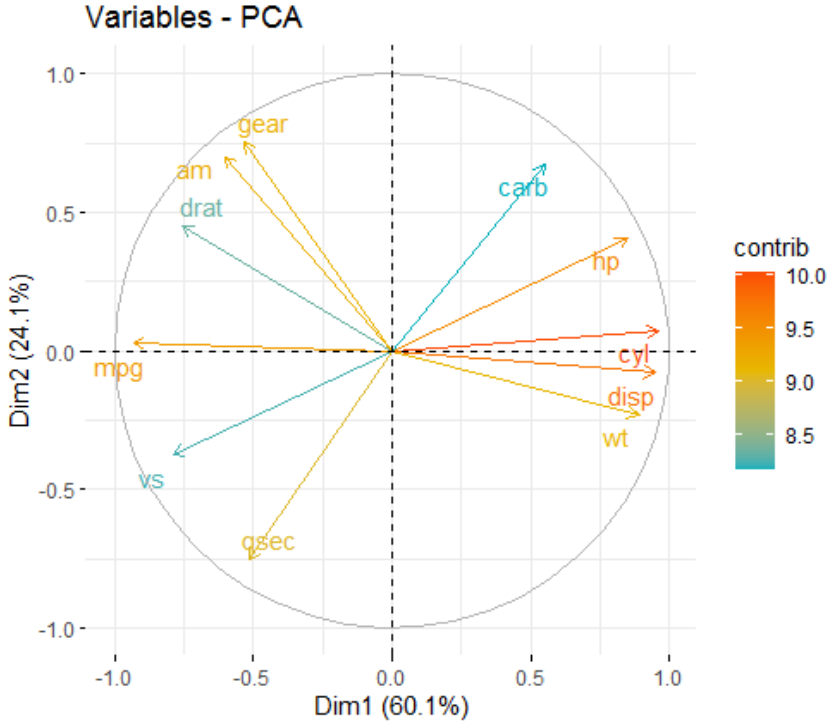



```{r}
library(factoextra)
Visual_PCA_Cov <- prcomp(mtcars, scale = FALSE)
fviz_pca_ind(Visual_PCA_Cov, col.ind = "cos2", gradient.cols =
c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE)

```



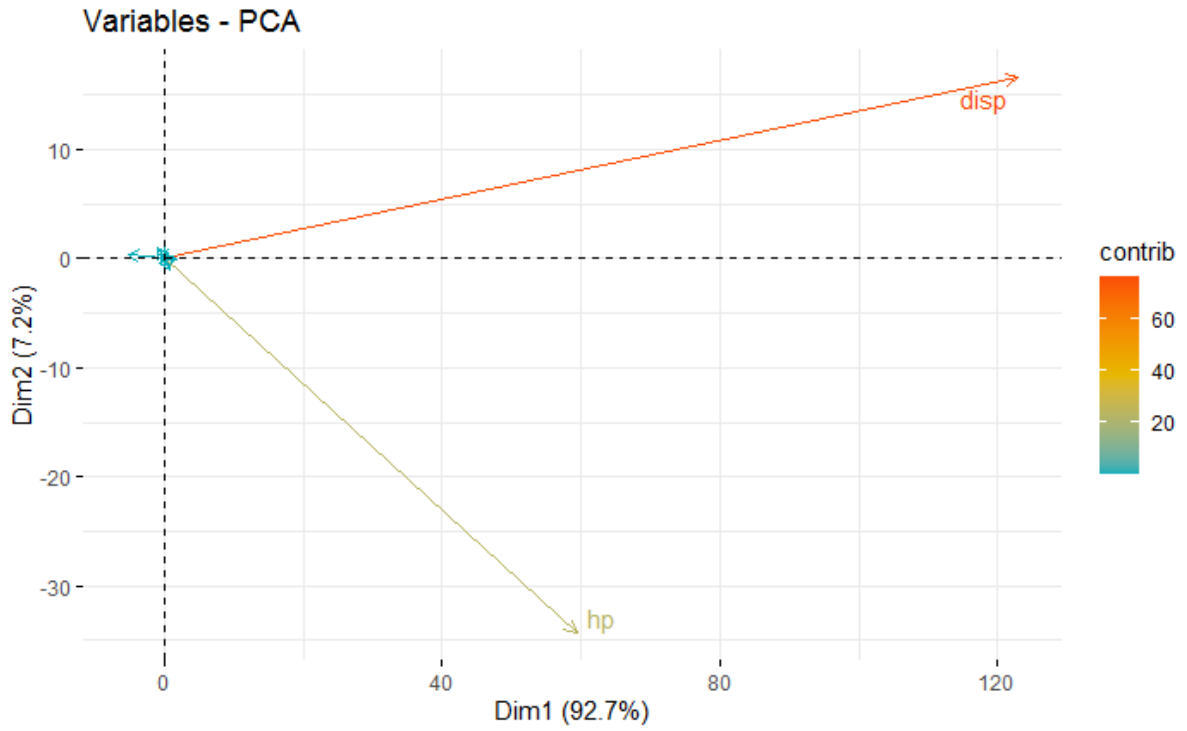
d. Plot di 2-dimensi untuk variabelnya

```{R}
#d Plot di 2-dimensi untuk variabelnya
fviz_pca_var(Visual_PCA_Cor, col.var = "contrib",
gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), repel =
TRUE)
```



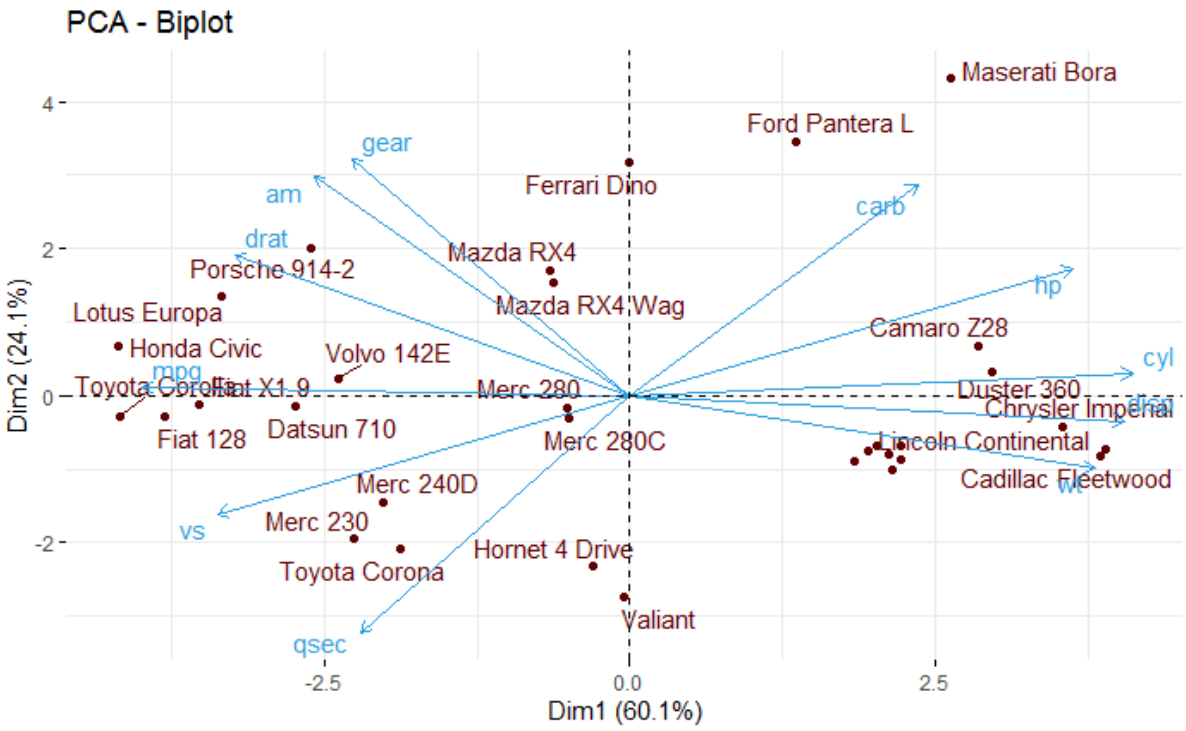

```{R}
```

```
fviz_pca_var(Visual_PCA_Cov, col.var = "contrib",
gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), repel =
TRUE)
```

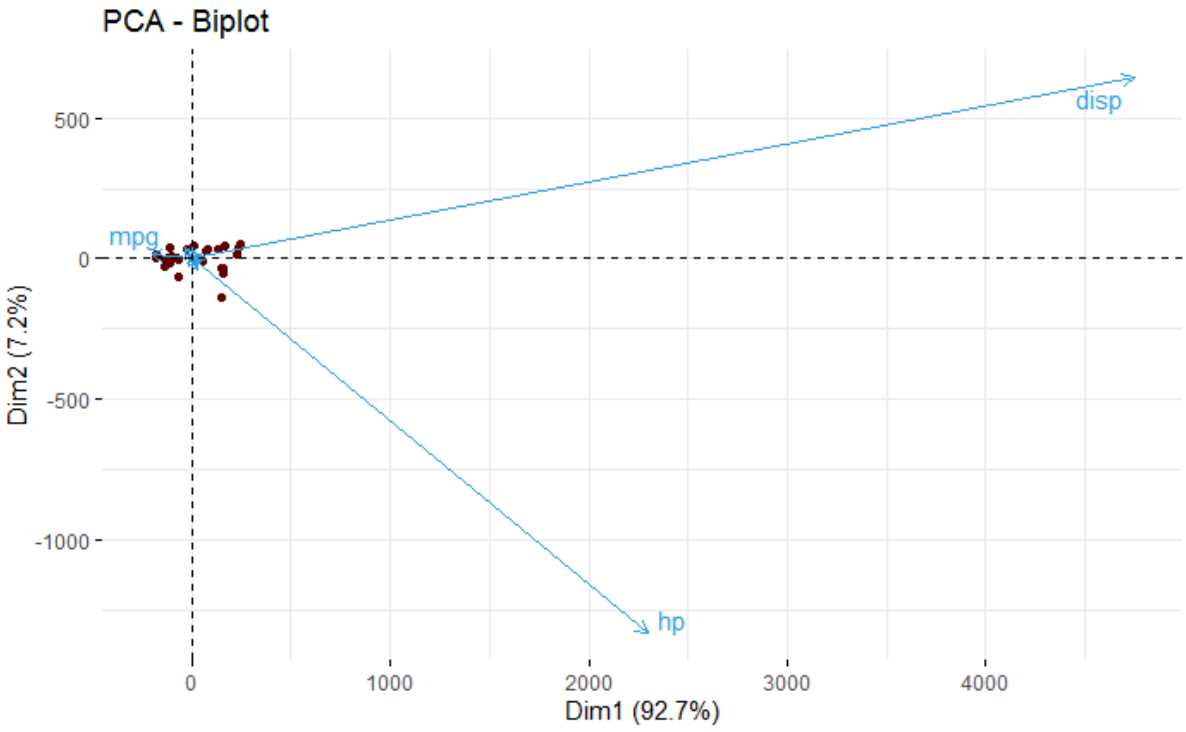


e. Bi-Plot di 2-dimensi untuk individu dan variabelnya

```{R}
#e.  Bi-Plot di 2-dimensi untuk individu dan variabelnya
fviz_pca_biplot(Visual_PCA_Cor, repel = TRUE, col.var =
"#2E9FDF", col.ind = "#600000")
```



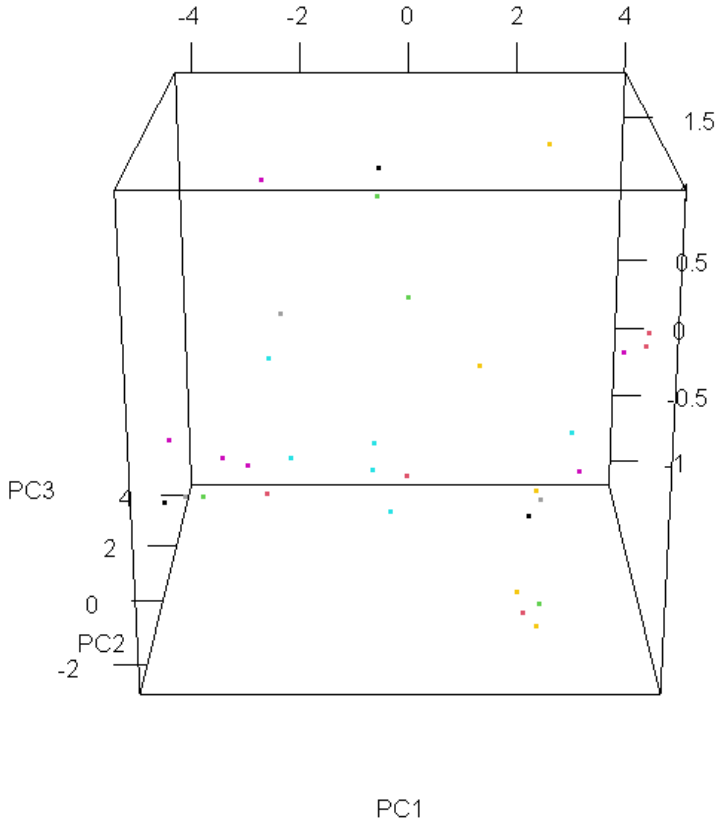

```{R}
fviz_pca_biplot(Visual_PCA_Cov, repel = TRUE, col.var = "#2E9FDF", col.ind = "#600000")
```



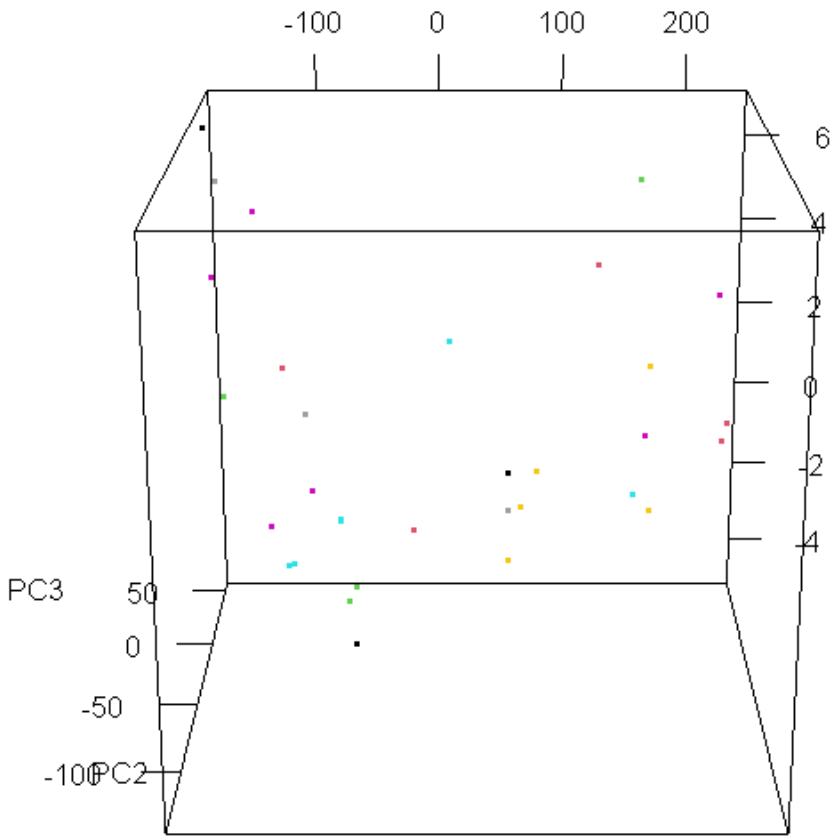
f. Plot PC di 3-dimensi

```{R}
#f. Plot PC di 3-dimensi
library(rgl)
```

```
plot3d(Visual_PCA_Cor$x, col = PCA_Data$mpg)
```




```{R}
library(rgl)
plot3d(Visual_PCA_Cov$x, col = PCA_Data$mpg)
```


-100
0
100
200
6
4
2
0
-2
-4
PC3
50
0
-50
-100
PC2
PC1

2. Carilah data yang sesuai untuk PCA dengan correlation

```{R}
#No 2
PCA_Standardized <- scale(x = PCA_Data)
Correlation_Matrix <- cor(PCA_Standardized)
Nilai_Eigen_Cor <- eigen(Correlation_Matrix)
Nilai_Eigen_Cor
PCA_Analisis_Cor <- prcomp(x = PCA_Data, scale. = TRUE, center
= TRUE)
PCA_Analisis_Cor
```

Description: df[,8] [6 x 8]

	mpg <dbl>	cyl <dbl>	disp <dbl>	hp <dbl>	drat <dbl>	wt <dbl>	qsec <dbl>	gear <dbl>
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	4
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	3
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	3
Valiant	18.1	6	225	105	2.76	3.460	20.22	3

6 rows

a. Tampilkan eigenvalue-nya (variansnya) → table dan grafik

```{R}
#a.  Tampilkan eigenvalue-nya (variansnya)
Nilai_Eigen_Cor <- eigen(Correlation_Matrix)
Nilai_Eigen_Cor
```

```
eigen() decomposition
$values
[1] 5.65593947 2.08210029 0.50421482 0.26502753 0.18315864 0.12379319 0.10506192 0.05851375
[9] 0.02219038

$vectors
           [,1]        [,2]        [,3]         [,4]       [,5]        [,6]        [,7]
 [1,]  0.3931477 -0.02753861 -0.22119309 -0.006126378  0.3207620 -0.72015586  0.38138068
 [2,] -0.4025537 -0.01570975 -0.25231615  0.040700251 -0.1171397 -0.22432550  0.15893251
 [3,] -0.3973528  0.08888469 -0.07825139  0.339493732  0.4867849  0.01967516  0.18233095
 [4,] -0.3670814 -0.26941371 -0.01721159  0.068300993  0.2947317 -0.35394225 -0.69620751
 [5,]  0.3118165 -0.34165268  0.14995507  0.845658485 -0.1619259  0.01536794 -0.04767957
 [6,] -0.3734771  0.17194306  0.45373418  0.191260029  0.1874822  0.08377237  0.42777608
 [7,]  0.2243508  0.48404435  0.62812782 -0.030329127  0.1482495 -0.25752940 -0.27622581
 [8,]  0.2094749 -0.55078264  0.20658376 -0.282381831  0.5624860  0.32298239  0.08555707
 [9,] -0.2445807 -0.48431310  0.46412069 -0.214492216 -0.3997820 -0.35706914  0.20604210
           [,8]        [,9]
 [1,]  0.12465987 -0.11492862
 [2,] -0.81032177 -0.16266295
 [3,]  0.06416707  0.66190812
 [4,]  0.16573993 -0.25177306
 [5,] -0.13505066 -0.03809096
 [6,]  0.19839375 -0.56918844
 [7,] -0.35613350  0.16873731
 [8,] -0.31636479 -0.04719694
 [9,]  0.10832772  0.32045892
```

b. Tampilkan matriks PC-nya

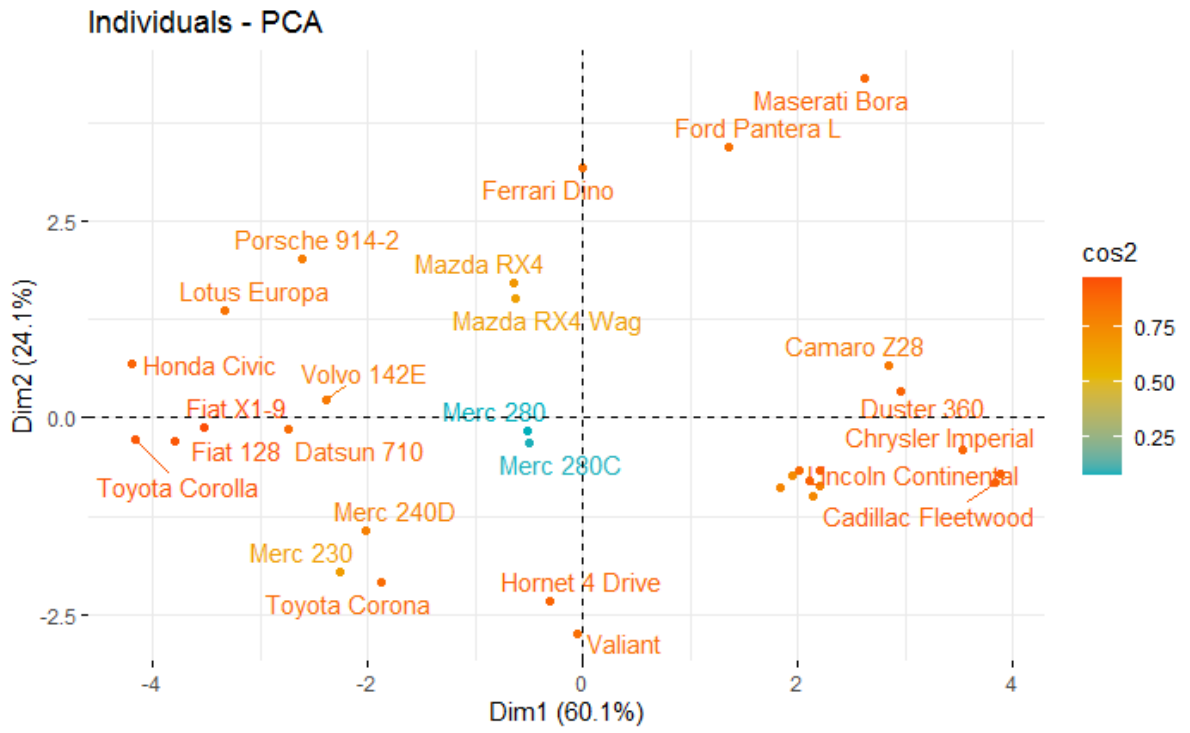
```{R}
#b.  Tampilkan matriks PC-nya
```

```
Correlation_Matrix <- cor(PCA_Standardized)
Correlation_Matrix
```

```
             mpg        cyl       disp         hp        drat         wt        qsec       gear
mpg   1.0000000 -0.8521620 -0.8475514 -0.7761684  0.68117191 -0.8676594  0.41868403  0.4802848
cyl  -0.8521620  1.0000000  0.9020329  0.8324475 -0.69993811  0.7824958 -0.59124207 -0.4926866
disp -0.8475514  0.9020329  1.0000000  0.7909486 -0.71021393  0.8879799 -0.43369788 -0.5555692
hp   -0.7761684  0.8324475  0.7909486  1.0000000 -0.44875912  0.6587479 -0.70822339 -0.1257043
drat  0.6811719 -0.6999381 -0.7102139 -0.4487591  1.00000000 -0.7124406  0.09120476  0.6996101
wt   -0.8676594  0.7824958  0.8879799  0.6587479 -0.71244065  1.0000000 -0.17471588 -0.5832870
qsec  0.4186840 -0.5912421 -0.4336979 -0.7082234  0.09120476 -0.1747159  1.00000000 -0.2126822
gear  0.4802848 -0.4926866 -0.5555692 -0.1257043  0.69961013 -0.5832870 -0.21268223  1.0000000
carb -0.5509251  0.5269883  0.3949769  0.7498125 -0.09078980  0.4276059 -0.65624923  0.2740728
           carb
mpg  -0.5509251
cyl   0.5269883
disp  0.3949769
hp    0.7498125
drat -0.0907898
wt    0.4276059
qsec -0.6562492
gear  0.2740728
carb  1.0000000
```

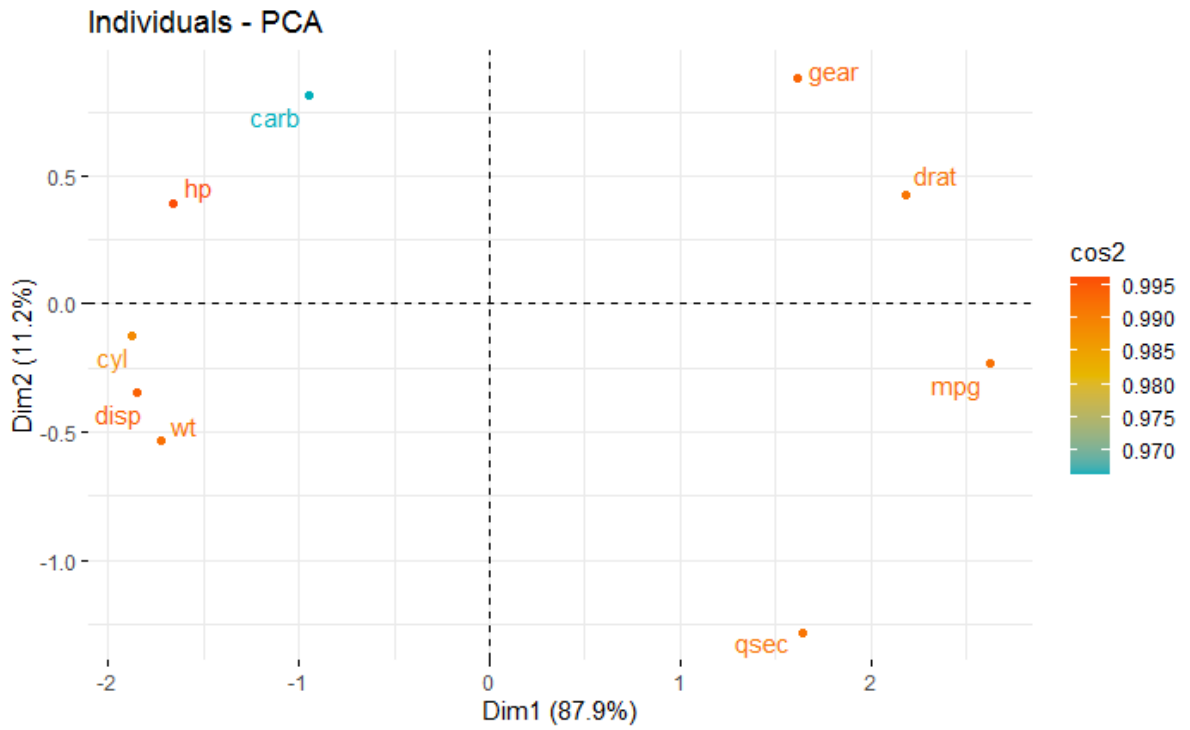
c. Plot di 2-dimensi untuk individunya

```{R}
#c.  Plot di 2-dimensi untuk individunya
library(factoextra)
Visual_PCA_Cor <- prcomp(Correlation_Matrix, scale = TRUE)
fviz_pca_ind(visual.pca.cor, col.ind = "cos2", gradient.cols =
c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE)
```

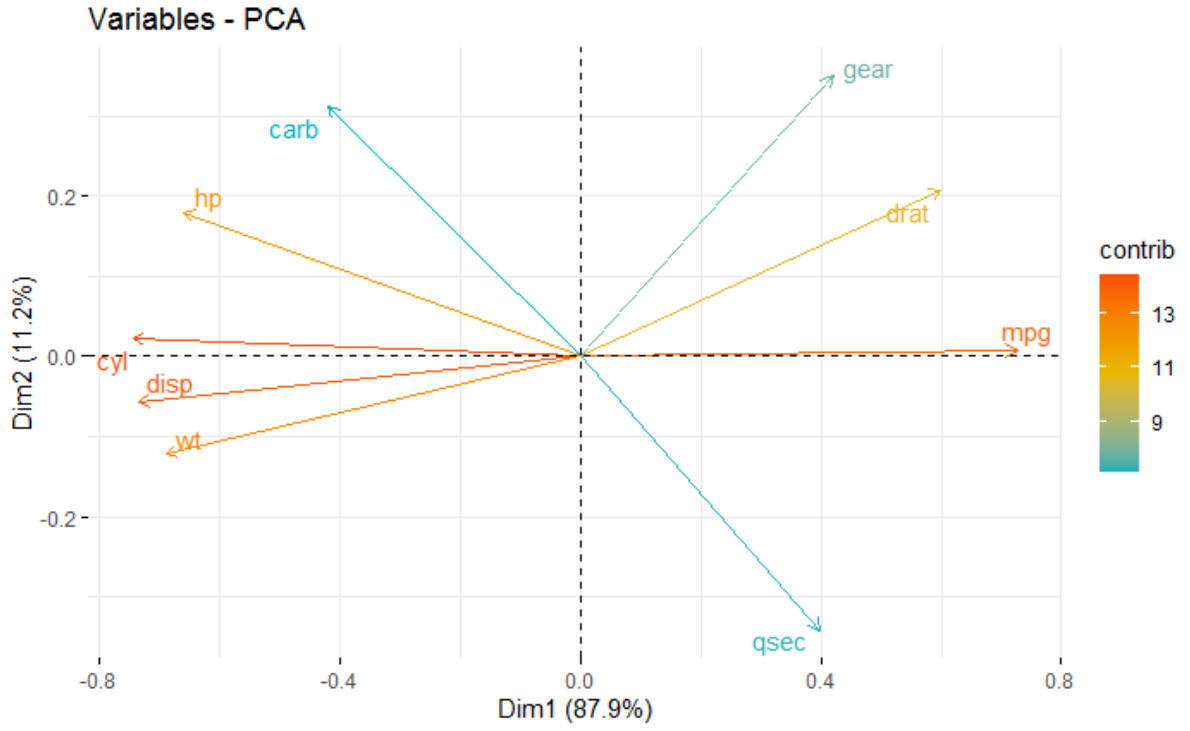



```{R}
library(factoextra)
Visual_PCA_Cor <- princomp(Correlation_Matrix, scale = TRUE)
```

```
fviz_pca_ind(Visual_PCA_Cor, col.ind = "cos2", gradient.cols =
c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE)
```

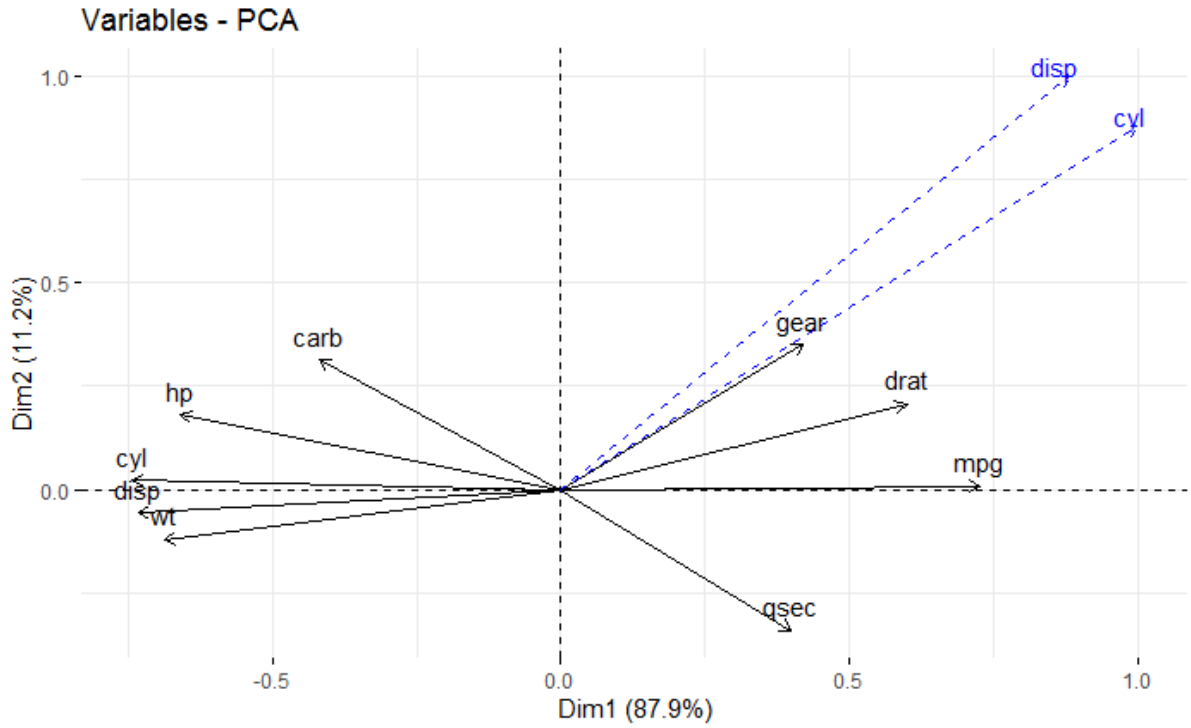



```{R}
fviz_pca_var(Visual_PCA_Cor, col.var = "contrib",
gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), repel =
TRUE)
```



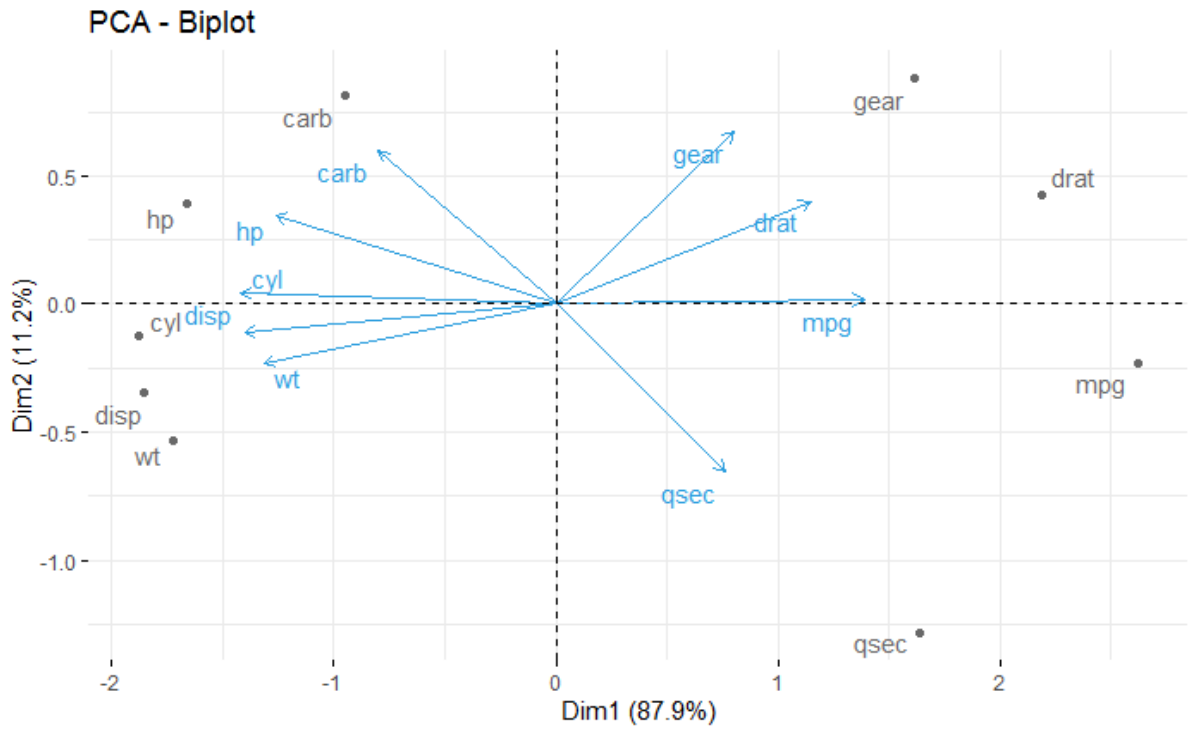
d. Plot di 2-dimensi untuk variabelnya

```{R}
#d.  Plot di 2-dimensi untuk variabelnya
PCA_Cor <- PCA_Data[1:23, 2:3, drop = FALSE]
PCA_Cor_New <- cor(PCA_Cor)
Correlation <- fviz_pca_var(Visual_PCA_Cor)
fviz_add(Correlation, PCA_Cor_New, color = "blue", geom =
"arrow")
```



e. Bi-Plot di 2-dimensi untuk individu dan variabelnya

```{R}
#e.  Bi-Plot di 2-dimensi untuk individu dan variabelnya
fviz_pca_biplot(Visual_PCA_Cor, repel = TRUE, col.var =
"#2E9FDF", col.ind = "#696969")
```



f. Plot PC di 3-dimensi

```{R}
#f.  Plot PC di 3-dimensi
library(rgl)
plot3d(Visual_PCA_Cor$x, col = PCA_Data$gear)
```